ABSTRACT
        Methods of regression commonality analysis are
generalized for use in canonical correlation analysis. An actual data
set (involving educators' attitudes toward death and age, locus of
control, religion, and occupational role in working with terminally
ill children) is employed to illustrate the extension. The method can
be applied with respect to each canonical function in an analysis to
determine the proportion of explanatory power of a variable set which
is unique, as well as the proportion which is shared with or common
to other variables also. (Author/GDC)

A MULTIVARIATE METHOD OF COMMONALITY ANALYSIS

Bruce Thompson   James H. Miller

University of New Orleans 70148

ABSTRACT

Methods of regression commonality analysis are generalized for use in canonical correlation analysis. An actual data set is employed to illustrate the extension. The method can be applied with respect to each canonical function in an analysis to determine the proportion of explanatory power of a variable or variable set which is unique and the proportion which is shared with or common to other variables also.

Canonical correlation analysis (Hotelling, 1935) has been available to researchers in theory for some 50 years. Kerlinger (1973, p. 652) suggests that "some research problems almost demand canonical analysis" and Cooley and Lohnes (1971, p. 176) argue that "it is the simplest model that can begin to do justice to this difficult problem of scientific generalization." Krus, Reynolds, and Krus (1976, p. 725) noted that, "Dormant for nearly half a century, Hotelling's (1935, 1936) canonical variate analysis has come of age. The principal reason behind its resurrection was its computerization and inclusion in major statistical packages."

Notwithstanding Levine's (1977, p. 8) assertion that "especially with respect to canonical correlation, there seem to be relatively few remaining puzzles to be solved," several puzzles involving the technique have been recognized and resolved in recent years. Noteworthy examples include specialized canonical rotation procedures (Bentler & Huba, 1982), backward variable elimination procedures (Thompson, 1982b), and procedures to estimate result invariance (Thompson, 1982a). Thompson (1984a) summarizes these and other recent extensions of canonical methods.

The present paper discusses methods of partitioning the ability of the variables in a canonical set to explain the variables in the other canonical set. These methods were originally developed for application in multiple regression analysis and have come to be referred to as "commonality analysis" (Pedhazur, 1982, pp. 199-211). Cooley and Lohnes (1976, p. 219) suggest that "the commonality method of partitioning of variance in multivariate regression is an informative, conservatively safe method for most situations." Lohnes and Cooley (1978, p. 17) argue that the results of commonality analyses "can well be the frosting on the cake that get the first attention." The fact that "virtually all of the commonly encountered parametric tests of significance can be

4

treated as special cases of canonical correlation analysis" (Knapp, 1978, p. 410) suggests that commonality analysis might also be applied usefully in the canonical case, if a canonical extension of this regression method can be specified. Data from an actual study will be used to illustrate the applications of this extension.

Commonality Analysis

Commonality analysis was originally developed for use in regression studies under the rubrics, "element analysis" (Newton & Spurrell, 1967) and "components analysis" (Mayeske, Wisler, Beaton, Weinfield, Cohen, Okada, Proshek & Tabler, 1969). The analysis indicates how much of the explanatory power of a variable is "unique" to the variable, and how much of the variable's explanatory ability is "common" to or also available from one or more other variables. Mood (1969) presents an algebraic rule for computing these variance partitions for any number of variables. In addition to tabling the computational procedures for studies involving up to five variables in a set, Seibold and McPhee (1979) provide a brief and understandable introduction to this method as applied in the regression case.

Two features of commonality analysis merit some explanation. First, commonalities should not be confused with the interaction effects investigated in ANOVA studies (Seibold & McPhee, 1979, p. 365). As Thompson (1984b, p. 8) explains, "interaction is the unique effect of two or more independent variables which in combination affect the dependent variable. Commonality indicates the proportion of predictive ability of a single variable that also happens to reside in another single predictor variable too; no unique effect of the predictors acting in combination is involved."

Second, although uniqueness estimates can never be negative, commonality partitions can be. This is counterintuitive since the result seems to indicate that two or more variables have in common the ability to explain less than zero percent of the variance in the criterion variable or variable set. Instead, negative commonalities frequently indicate the presence of suppressor effects (DeVito, 1976, p. 12). Beaton (1973, p. 12) provides a conceptual illustration of how a negative commonality can have important substantive implications:

> Both weight and speed are important to success as a professional football player and each would be moderately correlated with a measure of success in football. Weight and speed are presumably negatively correlated and would have a negative commonality in predicting success in football. If both weight and speed are known, one would expect to make a much better prediction of success using both variables to select fast, heavy men rather than just selecting the fastest regardless of speed. Thus the negative commonality indicates that explanatory power of either is greater when the other is used.

Beaton's conceptual example is supplemented by Seibold and McPhee's (1979, pp. 364-365) report of regression results from an actual cancer study that may well have been grossly misinterpreted if a commonality analysis detecting suppressor effects had not been conducted.

<center>Example Canonical Extension</center>

As noted previously, commonality analysis has not been applied to partition the squared canonical correlation coefficient, although commonality analysis has been applied to partition both squared multiple correlation and redundancy coefficients (Cooley & Lohnes, 1976). The extension proposed here capitalizes on the fact that canonical analysis can be performed using a regression computer routine if one set of variables is aggregated into canonical variate scores. Then the previously discussed rules for variance partitioning can also be applied in the multivariate case. An example application illustrates the procedure.

Miller (1984) reports a study of the death anxiety of various groups of educators who differ in their degree of contact with terminally ill youngsters. The theoretical framework and additional details of the study are presented in the original report, although a commonality analysis was not conducted by Miller. The tabled results are reported here in some detail for readers who wish to verify selected results.

<u>INSERT TABLE 1 ABOUT HERE.</u>

The first step in canonical commonality analysis involves the calculation in the usual manner of canonical functions and canonical correlation coefficients. These analyses are based on the correlation matrix presented in Table 1 and involved four criterion variables ("W" to "Z") and six predictor variables ("A" to "F"). The first canonical function is also presented in Table 1. The same procedures can be applied to partition the squared canonical correlation for any function, so discussion of only the first function should suffice as an example. The squared canonical correlation associated with the first function was .13739 (R_c = .37066, chi squared = 38.71, $\underline{df}$ = 24, $\underline{p}$ < .05).

The second step in canonical commonality analysis involves the calculation of canonical variate scores for all subjects. In the present example the partition was conducted to explore the ability of the six predictors to explain variance in the criterion variable set, so variate scores were computed for all 160 subjects using the function coefficients presented in Table 1. The function coefficients are multiplied times subjects' Z-scores on the criterion variables. For example, for the first subject the computations were:

$(.56672 \times \text{ZCOPING}) + (.63526 \times \text{ZFEAR}) + (-.49114 \times \text{ZSHORT}) + (.18460 \times \text{ZPAINFUL})$

$(.56672 \times .11606) + (.63526 \times -.17941) + (-.49114 \times -.31146) + (.18460 \times 1.00038)$

$$.06583 - .11397 + .15297 + .18467 = .28950$$

Of course, some statistics packages readily compute variate scores as an option. Otherwise with most packages the scores are readily calculated using automated compute procedures.

The third step involves the calculation of regression equations predicting the variate scores with all possible combinations of the predictor variable sets of interest. In the present study interest centered on the prediction of the four criterion variables (represented by the variate scores labelled "Y*" in Table 1) using the predictors age, locus of control, religious preference, and the three job role contrasts as a set (labelled "D*"). The 15 combinations of these predictors, their predictive abilities, and the use of the values in computing canonical uniqueness and commonality estimates are presented in Table 2. Again, readers can verify these results by analyzing the relevant portion of the Table 1 correlation matrix, this time involving variables labelled "Y*", and "A" through "F".

<u>INSERT TABLE 2 ABOUT HERE.</u>

8

It should be noted that coefficients involving the "D*" variable set actually involved three predictors, "D" through "F". For example, the squared multiple correlation involving the canonical variate scores ("Y*") and the three predictors in the job role set ("D*") was .01820. Thus, the example makes the point that these procedures can be employed to examine either individual variables or various sets of variables. Of course, the squared multiple correlation using all the predictors to predict "Y*" (.13739) exactly equals the squared canonical correlation since in the full model case the regression analysis is a canonical analysis.

Table 3 summarizes the results in the manner recommended for the regression case by Mayesks et al. (1969). As should be expected, since the 15 values are partitions of the squared canonical correlation, the 15 partitions, sum to the value of the squared canonical correlation. Furthermore, as noted in the table, the analysis also partitions the individual predictive power of the predictor variables represented by the squared correlations with the canonical variate scores.

INSERT TABLE 3 ABOUT HERE.

Discussion

Canonical commonality analysis is not without limits. For example, as Newton and Spurrell (1967, p. 61) note, "it is difficult to see that statistical theory will be able to give sampling errors which can be used in meaningful tests for secondary elements [commonalities] since they are obviously not independent statistical quantities." However, the inability to statistically test commonalities does not seem inherently debilitating since the technique's focus is on interpretation after a significant canonical correlation has already been detected. Furthermore, the focus is

consistent with the recognition that statistical significance is primarily a function of sample size (Carver, 1978), and with an emphasis on effect sizes in meta-analysis and other applications (Glass, McGaw & Smith, 1981).

Commonality analysis honors the relationships among variables in a set, and as Seibold and McPhee (1979, p. 355) argue:

> Advancement of theory and the useful application of research findings depend not only on establishing that a relationship exists among predictors and the criterion, but also upon determining the extent to which those independent variables, singly and in all possible combinations, share variance with the dependent variable. Only then can we fully know the relative importance of independent variables with regard to the dependent variable in question.

As Mood (1969, p. 480) notes, "The independent variables in any social process, and certainly in education, are highly correlated among themselves, and this kind of partition of variance provides measures of the extent to which they overlap each other in their association with the dependent variable."

A final benefit of canonical commonality analysis is didactic. The extension of a regression technique to the canonical case reinforces the recognition (Knapp, 1978) that canonical analysis is the most general case of parametric significance testing. Understanding the linkages among various statistical methods may facilitate more considered analytic choices in contemporary research practice.

References

Beaton, A.E. (1973). Commonality. Unpublished manuscript, Educational Testing
Service, Princeton, NJ.

Bentler, P.M., & Huba, G.J. (1982). Symmetric and asymmetric rotations in
canonical correlation analysis: New methods with drug variable examples. In
N. Hirschberg & L.G. Humphreys (Eds.), Multivariate applications in the
social sciences. Hillsdale, NJ: Erlbaum.

Carver, R.P. (1978). The case against statistical significance testing. Harvard
Educational Review, 48, 378-399.

Cooley, W.W., & Lohnes, P.R. (1971). Multivariate data analysis. New York:
Wiley.

Cooley, W.W., & Lohnes, P.R. (1976). Evaluation research in education. New
York: Irvington Publishers.

DeVito, P.J. (1976). The use of commonality analysis in educational research.
Paper presented at the annual meeting of the New England Educational Research
Association, Provincetown, Massachusetts. (ERIC Document Reproduction
Service No. ED 146 218)

Glass, G.V, McGaw, B., & Smith, M.L. (1981). Meta-analysis in social research.
Beverly Hills, CA: SAGE.

Hotelling, H. (1935). The most predictable criterion. Journal of Experimental
Psychology, 26, 139-142.

Kerlinger, F.N. (1973). Foundations of behavioral research (2nd ed.). New York:
Holt, Rinehart & Winston.

Knapp, T.R. (1978). Canonical correlation analysis: A general parametric
significance testing system. Psychological Bulletin, 85, 410-416.

Krus, D,J., Reynolds, T.S., & Krus, P.H. (1976). Rotation in canonical variate

analysis. _Educational and Psychological Measurement, 36_, 725-730.

Levine, M.S. (1977). _Canonical analysis and factor comparison._ Beverly Hills, CA: SAGE.

Lohnes, P.R., & Cooley, W.W. (1978). _Regarding criticisms of commonality analysis._ Paper presented at the annual meeting of the American Educational Research Association, Toronto.

Mayeske, G.W., Wisler, C.E., Beaton, A.E., Jr., Weinfield, F.D., Cohen, W.M., Okada, T., Proshek, J.M., & Tabler, K.A. (1969). _A study of our nation's schools._ Washington, D.C.: Office of Education, U.S. Department of Health, Education and Welfare.

Miller, J.H. (1984). _Demographic, personality and experential correlates of death anxiety of educators._ Unpublished doctoral dissertation, University of New Orleans, New Orleans.

Mood, A.M. (1969). Macro-analysis of the American educational system. _Operations Research, 17_, 770-784.

Newton, R.G., & Spurrell, D.J. (1967). A development of multiple regression for the analysis of routine data. _Applied Statistics, 16_, 51-64.

Pedhazur, E.J. (1982). _Multiple regression in behavioral research_ (2nd ed.). New York: Holt, Rinehart and Winston.

Seibold, D.R., & McPhee, R.D. (1979). Commonality analysis: a method for decomposing explained variance in multiple regression analyses. _Human Communication Research, 5_, 355-365.

Thompson, B. (1982a). _Comparison of two methods for computing canonical invariance coefficients._ Paper presented at the annual meeting of the Southwest Educational Research Association, Austin.

Thompson, B. (1982b). A logic for stepwise canonical correlation analysis.

Perceptual and Motor Skills, 54, 879-882.

Thompson, B. (1984a). Canonical correlation analysis: Uses and interpretation. Beverly Hills, SAGE.

Thompson, B. (1984b). Coding and commonality analysis: Non-ANOVA methods for analyzing data from experiments. Paper presented at the annual meeting of the Mid-South Educational Research Association, New Orleans.

Table 1

Correlation and Canonical Function Coefficients

Criterion Variables	W	X	Y	Z	Y*	A	B	C	D	E	Function Coefs.
Coping with Death (W)											.567
Fear of Dying (X)	-00003										.635
Shortness of Life (Y)	-00001	00000									-.491
Painful Death (Z)	00000	00000	00001								.185
Predictor Variables											
Age (A)	-15690	-17885	13266	-09564	-28534						-.698
Locus of Control (B)	12944	20089	-09838	03795	25630	-19567					.610
Religious Preference (C)	08211	03595	-09457	-18293	08205	-15454	-08741				.227
Role Group Contrast 1 (D)	-09408	06045	04350	-03547	-04283	-11957	07962	02673			-.254
Contrast 2 (E)	06376	13514	05524	06344	10657	-35704	12548	04629	00000		-.049
Contrast 3 (F)	13771	-01418	-04437	-10881	07074	-26963	07028	39279	00000	00000	-.129

Note: "Y*" represents the set of canonical variate scores for the criterion variable set.

15

14

Table 2
Canonical Commonality Computations

Unique to A R^2 $- R^2$
            A,B,C,D*   B,C,D*
            .13739 $-$ .08600 = .05139

Unique to B R^2 $- R^2$
            A,B,C,D*   A,C,D*
            .13739 $-$ .08959 = .04780

Unique to C R^2 $- R^2$
            A,B,C,D*   A,B,D*
            .13739 $-$ .13157 = .00582

Unique to D* R^2 $- R^2$
            A,B,C,D*   A,B,C
            .13739 $-$ .12727 = .01012

Common to A,B R^2 $+ R^2$ $- R^2$ $- R^2$
            A,C,D*   B,C,D*   C,D*   A,B,C,D*
            .08959 + .08600 $-$ .02122 $-$ .13739 = .01698

Common to A,C R^2 $+ R^2$ $- R^2$ $- R^2$
            A,B,D*   B,C,D*   B,D*   A,B,C,D*
            .13157 + .08600 $-$ .07807 $-$ .13739 = .00211

Common to A,D* R^2 $+ R^2$ $- R^2$ $- R^2$
            A,B,C   B,C,D*   B,C   A,B,C,D*
            .12727 + .08600 $-$ .07668 $-$ .13739 = $-$.00080

Common to B,C R^2 $+ R^2$ $- R^2$ $- R^2$
            A,B,D*   A,C,D*   A,D*   A,B,C,D*
            .13157 + .08959 $-$ .08751 $-$ .13739 = $-$.00374

Common to B,D* R^2 $+ R^2$ $- R^2$ $- R^2$
            A,B,C   A,C,D*   A,C   A,B,C,D*
            .12727 + .08959 $-$ .08290 $-$ .13739 = $-$.00343

Common to C,D* R^2 $+ R^2$ $- R^2$ $- R^2$
            A,B,C   A,B,D*   A,B   A,B,C,D*
            .12727 + .13157 $-$ .12321 $-$ .13739 = .00941

Common to A,B,C R^2 $+ R^2$ $+ R^2$ $+ R^2$ $- R^2$ $- R^2$ $- R^2$ $- R^2$
            C,D*   B,D*   A,D*   A,B,C,D*   D*   B,C,D*   A,C,D*   A,B,D*
            .02122 + .07907 + .08751 + .13739 $-$ .01820 $-$ .08600 $-$ .08959 $-$ .13157 = $-$.00117

Common to A,B,D* R^2 $+ R^2$ $+ R^2$ $+ R^2$ $- R^2$ $- R^2$ $- R^2$ $- R^2$
            C,D*   B,D*   A,C   A,B,C,D*   C   B,C,D*   A,C,D*   A,B,C
            .02122 + .07807 + .08290 + .13739 $-$ .00673 $-$ .08600 $-$ .08959 $-$ .12727 = .00999

Common to A,C,D* R^2 $+ R^2$ $+ R^2$ $+ R^2$ $- R^2$ $- R^2$ $- R^2$ $- R^2$
            B,D*   B,C   A,B   A,B,C,D*   B   B,C,D*   A,B,D*   A,B,C
            .07807 + .07668 + .12321 + .13739 $-$ .06569 $-$ .08600 $-$ .13157 $-$ .12727 = $-$.00635

Common to B,C,D* R^2 $+ R^2$ $+ R^2$ $+ R^2$ $- R^2$ $- R^2$ $- R^2$ $- R^2$
            A,D*   A,C   A,B   A,B,C,D*   A   A,C,D*   A,B,D*   A,B,C
            .08751 + .08290 + .12321 + .13739 $-$ .08142 $-$ .08959 $-$ .13157 $-$ .12727 = $-$.01001

Common to all R^2 $+ R^2$ $+ R^2$ $+ R^2$ $+ R^2$ $+ R^2$ $+ R^2$ $+ R^2$
            D*   C   B   A   B,C,D*   A,C,D*   A,B,D*   A,B,C
            .01820 + .00673 + .06569 + .08142 + .08600 + .08959 + .13157 + .12727
       $- R^2$ $- R^2$ $- R^2$ $- R^2$ $- R^2$ $- R^2$ $- R^2$
         C,D*   B,D*   B,C   A,D*   A,C   A,B   A,B,C,D*
       $-$ .02122 $-$ .07807 $-$ .07668 $-$ .08751 $-$ .08290 $-$ .12321 $-$ .13739   = .01066

16

Table 3

Summary of Commonality Results

	A	B	C	D*
Unique to Age (A)	5.1%			
Unique to Locus of Control (B)		4.8%		
Unique to Religious Preference (C)			.6%	
Unique to Role Contrasts (D*)				1.0%
Common to A,B	1.7%	1.7%		
Common to A,C	.2%		.2%	
Common to A,D*	-.1%			-.1%
Common to B,C		-.4%	-.4%	
Common to B,D*		-.3%		-.3%
Common to C,D*			.9%	.9%
Common to A,B,C	-.1%	-.1%	-.1%	
Common to A,B,D*	1.0%	1.0%		1.0%
Common to A,C,D*	-.6%		-.6%	-.6%
Common to B,C,D*		-1.0%	-1.0%	-1.0%
Common to A,B,C,D*	1.1%	1.1%	1.1%	1.1%
Sum of Partitions in Column	8.3%	6.7%	.7%	2.0%
Squared Correlation with Canonical Variate Scores	.0814	.0657	.0067	.0182

Note: The sum of the 15 partitions (.13878) equals the squared canonical correlation coefficient within rounding error.